

Web Statistics using Analog at NNDC

Presentation for the USNDP 2003 Meeting
November 6-7, 2003

Thomas W. Burrows



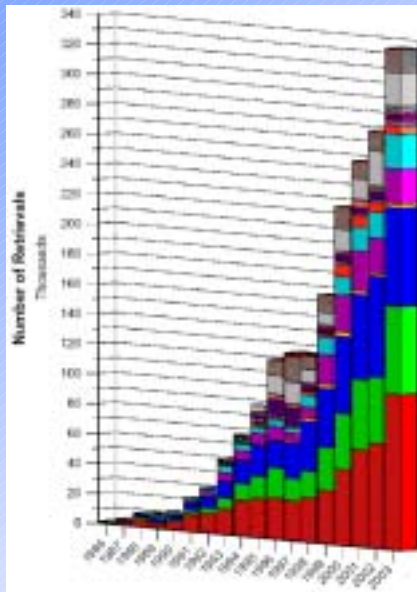
Web Statistics using Analog (NNDC) – 1

Summary

- Unable to reproduce statistics already generated.
- May be useful for certain general demographic studies although there are some limitations or problems one should be aware of.



Web Statistics using Analog (NNDC) – 2 History



Web Statistics using Analog (NNDC) – 3

NNDC Procedure

■ Non CGI Retrievals

- Certain status or a zero byte count returned excluded.
- Requests only for the top directory or navigation HTML excluded.
- “Decorative” or navigation icons excluded.
- Only specific subdirectories are considered. Dependent on the organization. only certain files are included or specific files are excluded.



Web Statistics using Analog (NNDC) – 4

NNDC Procedure (Cont.)

- Counting dependent on the organization of the information.
 - All contents of a subdirectory completely related: All requests are treated as only one retrieval unless the user moves to another NNDC page or more than one hour has elapsed between requests.
 - Contents of a subdirectory represent different discrete sets of data: All requests for one file are treated as only one retrieval unless the user requests a different file, moves to another NNDC page, or more than one hour has elapsed between requests.



Web Statistics using Analog (NNDC) – 4

NNDC Procedure (Cont.)

■ CGI Retrievals

- Only successful retrievals are logged to a special file.
- Retrieval criteria are also saved to this file.
- Seems to be similar to what can be done with asp or jsp, SQL, and modern relational databases.

■ robots.txt in place – Complete exclusion

- Disadvantage: NNDC under represented in search engines
- Advantage: only access is through “leakage” from other sites.



Web Statistics using Analog (NNDC) – 5

Analog Exercise

- Only www.nndc.bnl.gov 2002 access log used.
www2.nndc.bnl.gov provides access to:
 - NSR SQL Server
 - Mirror of the K.A.E.R.I. Table of Nuclides
- QuickDNS used in an attempt to resolve all numeric IP addresses for 2001 through 2002.
 - 1 to 3 minutes per month.
 - 6614 out of 24784 resolved (26.7%).



Web Statistics using Analog (NNDC) – 6

Analog Exercise (Cont.)

	Analog 5.31		NNDC Method
	Default	Configured	
Successful Retrievals	2,383,648 (49.47 Gbytes)	324,744 (38.89 Gbytes)	243,759
Codes		12,098	4,113
Wallet Cards		34,068	23,801
Capture γ's		13,441	13,296
NSR		35,147	40,515
NuDat		67,960	65,906



Web Statistics using Analog (NNDC) – 7

Analog Exercise (Cont.)

STATUSEXCLUDE 304

FILEALIAS /nndcscr/testwww/nsr*. * xxx.nsr

FILEALIAS /htbin/nsrlink.cgi* xxx.nsr

REQALIAS xxx.nsr [NSR-Retrievals]

FILEEXCLUDE /cgi-bin-tb/nsrlist*

FILEEXCLUDE /htbin/nsrlist*

FILEEXCLUDE /htbin/nsronl.cgi

FILEEXCLUDE /nndc/nsr/*

TYPEALIAS .nsr "NSR Retrievals - HTML, TXT, CGI"

FILEEXCLUDE /nndc/nudat/*.gif

FILEEXCLUDE /nndc/nudat/*

FILEEXCLUDE /nndcscr/testwww/nudat*

FILEALIAS /htbin/nudat.cgi* [Nudat]

